

Building a Lasting Coordination Group

(Advice from IETF)

Participants

- Rebecca Koskela
- Andrew Maffei
- Robert McDonald
- Mimi McClure
- Peter Wittenburg
- Martha Anderson
- Reagan Moore
- Ray Pelletier
- Russ Housley
- Beth Plale
- Jim Myers
- Ann Zimmerman

Topics

- IETF as Model for Coordination
 - What are the areas?
 - Are there approaches we can follow?
- Are we too late? Not too late, repositories already exist minimizing uncertainty.
 - Can propose ways to interact with existing data
 - Can propose new abstractions for discovery, access, process, manage
- Can we build on rough consensus on running code as way to make progress
 - De facto standards work if they are good enough
- Coordination of needs assessment
 - Interoperable use cases
 - Can each use case build on common infrastructure
 - Can your system support another group's use case
 - Common way of expressing a use case
 - Survey of users
 - Could provide as reports instead of running code?

Brain Storming

- Build high-level model to establish common vocabulary
- Be able to look at data object independently of domain
 - Example is use of DFDL or East to describe structure rather than required structure
- Within domain can apply domain standard to object
 - Domain specific semantics, formats, ontology, processes
- Coordinate use cases
 - Build standard document for describing use cases
 - Build representation of existing knowledge base, best practices, policies, procedures.
 - This implies best practices constitute a use case.
 - Demonstrate interoperability across use cases
- Interoperable services
 - Your service can access my repository
- This is our opportunity to create a culture

Areas

- Common architecture
 - Common language for describing architecture
 - OASIS? / OAI-PMH?
- Data description
 - Should this be a domain activity?
- Protocol level
 - Interoperability of services
- Policy / Procedure
 - How harmonize data sets for use
- User client
 - Features

Experience

- IETF evolved the terminology over time
 - Did agree on gross structure and layers
 - Should provide similar gross structure for data cyberinfrastructure
 - Open issues from first meeting are still open
 - IETF said no to communities that were too small or not generally useful
 - Need active management of process to control effort
 - Easier to involve vendors than administrators
 - Provide rationale for a given standard
- Have mature communities within data cyberinfrastructure?
 - Need to build upon the existing expertise
 - Can follow IETF in terms of comparing running code
 - Expect level of maturity is lower, implying the developer and administrator community are the same. When more mature, communities will separate
- How do we take lessons learned for globally useful infrastructure
 - What didn't work, what did work
 - IETF started as a communication mechanism for interoperability agreements.
 - Standards should be no better than have to be

Common Problems

- Common problems exist
 - Registries
 - How to make interoperable
 - How to get contributions
 - Library of Congress
 - Sharing through registries – have a dozen attempts
 - Need bulk load to keep current
 - Need to get more out than contribute
 - Want common machinery for implementing a registry
- Make policy explicit
 - Be able to characterize policies
- Making procedures explicit

Approaches

- Build community for exchanging experience
 - Best practices
- Or build community for creating interoperable infrastructure
 - Code
 - Standards

Challenges

- How to involve developers as well as users
 - Need to prioritize features and say no
 - What is the minimum number of users for development to go forward?
 - Depends upon importance. Can go forward if involves all experts within area.
- How track progress
 - Feature creep versus implementation effort
 - Need to compare with actual practice, and stop development when features do not have momentum
- Define culture for data cyberinfrastructure
 - Expectations for participation, how evolve infrastructure
 - Who is in charge? Can group consensus help manage evolution.
 - Do not expect a final solution
- When succeed with technical problem, there is a social implication
 - IETF does not track, but is important to data community
 - Comparison of user requirements does lead to generic software
 - Can we use projects as intermediary to gather user requirements?
- How manage volunteer community
 - How are we different?

Vision

- What is the vision?
 - Joining forces to stay ahead of the data deluge? (Plale)
 - Access by anyone anywhere (Pelletier)
- Example vision (Moore)
 - Storage based processing
 - Automated feature detection and indexing
 - Policy-based data management
 - Policy-encoded objects, policies are exchanged with govern data use
 - Manage provenance of processes
 - Replay workflows, enable reproducible science
 - Registration of workflows
 - Enable knowledge repositories
 - Associate community index with automated feature index, to enable cross correlation of features

How to Organize

- Do we have a self-supporting model for collaboration?
 - Membership fee
 - Vendor support
 - Do not lock into a model, but start thinking about how acquire administrative support
 - Is there an organization we can attach to?
- Or do we rely upon NSF funding to get started?
 - IETF had 10 years of NSF funding
 - NSF is interested in helping initiate activity, but encourages an independent society quicker than 25 years.
 - Is the Internet Society a viable support system?
 - EUDAT has funding from EU, and happy to collaborate
 - Will organize two workshops

Issues

- Take an approach that works for our culture
- Want to find groups that are further along
 - This accelerates work effort
- What other organizations should be involved?
 - Life Science Alliance
 - Want information about this organization
 - National Lambda Rail